NASA Web Strategy

White Paper: NASA's Future on the Web

DRAFT 1 – 4/12/2008
Chris C. Kemp

Executive Summary

NASA is comprised of a diverse portfolio of organizations unified by the common mission of expanding human knowledge and presence in the universe.

Over the past 50 years, NASA has taken people into space, to the surface of the moon, and has allowed people to see Earth, other planets in our solar system, and the depths of space through the eyes of satellites, telescopes, and robots, and through the cameras of astronauts. With each passing decade, advances in technology made images clearer, the information coming back from space richer, and the world smaller.

In 1969, America watched on television as NASA landed on the Moon. In 1997, America downloaded "live" images of Mars on the Internet. In 2018, how will America experience and participate in the new era of space exploration when NASA returns to the Moon and voyages out into the solar system? How will NASA leverage zettabytes of information to benefit humanity?

> "NASA's activities shall be conducted so as to contribute materially to … the expansion of human knowledge of the Earth and of phenomena in the atmosphere and space …"
>
> *Declaration of Policy and Purpose of NASA, Section 102 of the NASA Space Act*

This white paper presents an exciting new vision for how NASA will leverage the Internet over the next decade and outlines a practical plan which NASA can begin executing immediately that will dramatically increase America's participation in its space program, at a fraction of the cost of what NASA spends on its present public web sites.

Introduction

With its roots in the early days of spaceflight,[1] the Internet is increasingly providing an interface to every part of our daily lives on Earth. As we continue to expand humanity's presence in space, NASA must not forget that its charter requires that the information gained from exploration must be shared[2]. NASA should build a web platform through which the world will be able to experience, participate, and engage first hand our growing knowledge of and presence in the Solar System and beyond.

Where We Are Today

NASA's Public Web Presence[3] consists of a widely scattered footprint of several thousand static, text-heavy websites with www.nasa.gov at its center.

---

[1] DARPANET, the precursor to the modern Internet, was set up as a reaction to the former Soviet Union's launch of Sputnik.

[2] Section 203 "Functions of the Administration" (a) (3) of the NASA Space Act. "[NASA]… shall... provide for the widest practicable and appropriate dissemination of information concerning its activities and the results thereof"

[3] Public Web Presence as defined here consists of publically available (Internet) websites only. Intranet and extranet (password protected publically accessible sites) are excluded. In the future, these can be made available through secure web services.

Today, over 90% of the American public begins their search for information on the Internet using popular search portals such as Google, Yahoo!, or MSN. Most of NASA's Public Web Presence has been indexed by search engines such as these.
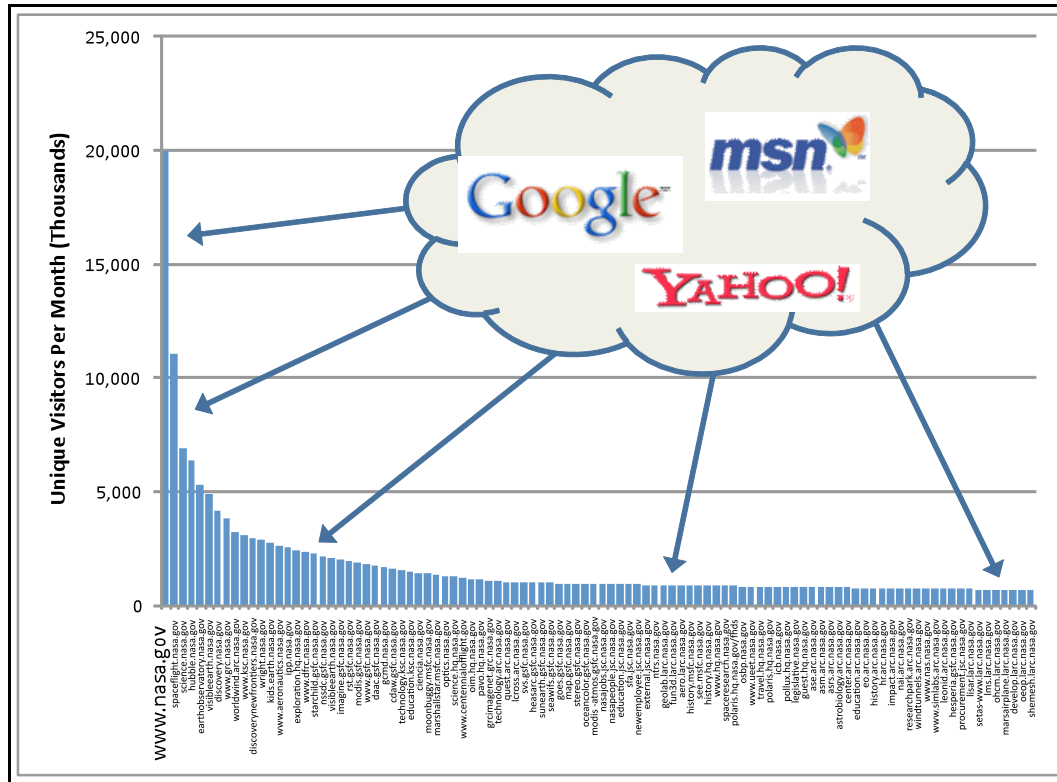

Figure 1. Traffic to NASA Public Web Sites[4].

Figure 1 illustrates this clearly. Search engines bring the public directly to the content they are looking for, which for reasons explained below, is often hosted on one of several thousand NASA websites, circumventing the www.nasa.gov site entirely. The total traffic to these sites far exceeds the traffic to www.nasa.gov. This presents many challenges, including:

1. **Cost**. This federated web architecture is incredibly expensive to maintain.

2. **Confusion**. It is very difficult to find information because there is no consistent information-architecture across these sites.

3. **Security**. Each of these thousands of independently managed websites presents a massive publically-accessible "attack surface" and constitutes a significant threat to the Agency's security posture. Thousands of independently-operated public-facing NASA websites are continuously being targeted, defaced, and compromised.

---

[4] These figure presents estimated, not actual monthly unique visitor statistics and is for illustrative purposes only.

4. **Self-competition**. NASA websites compete with each other for dominance in certain search engine rankings, often resulting in non-NASA sites being easier to find using space-related keywords.

These individual websites host a wide variety of content targeted to niche audiences. Each site is typically independently developed and maintained and hosted on independent physical servers across the Agency. With each content provider serving its own website, a variety of inconsistent infrastructures and design architectures have propagated. Development and content creation are largely outsourced to local contractors, resulting in duplicative efforts with little interaction between different teams. For all of these reasons, NASA's web presence is ineffective, inefficient, and costly to the Agency.

An effort led by the Public Affairs Office five years ago to address these challenges resulted in the creation of the "NASA Portal". While successfully consolidating some Public Affairs content onto www.nasa.gov, and consolidating about 75 websites into a common infrastructure, this effort resulted in the development of a proprietary and closed content management software environment that is largely inaccessible to anyone who is not a Public Affairs official. Further, the Portal exists in an expensive external hosting environment that is unaffordable and thus further inaccessible to most users. Additionally, the "NASA Portal" stifled innovation as changes required coordination with NASA Headquarters and procurement through the supplier of the "NASA Portal" contractors, developers, software, servers, networks, and infrastructure.

The software that powers www.nasa.gov today was built largely to facilitate Public Affairs Office editorial oversight of all content. The functionality available on the site is limited in scope, being focused around the News Magazine publishing model, making it unattractive to content owners and producers whose content does not comport with this model.

For these reasons, many content developers still decide against going through the trouble and expense of utilizing the "NASA Portal" and purchase their own servers, hire their own software developers and content creators, and set up duplicative infrastructure over which they have full ownership, control, and development flexibility. Today, over 2,000 publicly accessible websites are hosted by NASA, and after five years, less than 75 are hosted in the "NASA Portal."

How We Got Here

NASA was among the first to embrace the Internet as a medium for information dissemination. As a result, the number of publicly accessible websites on the NASA domain has grown from 100 in the early 1990s to well over 1,000 individually hosted websites in 2002.

Cognizant of the dispersion problem and the need for consolidation and organization, then NASA Administrator Sean O'Keefe initiated an effort in 2002 to begin work on restructuring NASA's web presence. The Public Affairs Office developed and implemented the "NASA Portal," which launched in 2003, with the specific purpose of catering to identified, yet amorphous, external audiences: the general public, students, teachers, children, and the mass media.

It was expected that the NASA Portal would consolidate NASA's Internet presence; however, despite subsidization and clear policies mandating that all websites be migrated to the Portal,

consolidation has yet to fully materialize. Five years after the Portal's original launch, all Centers and Mission Directorates still host numerous websites outside the www.nasa.gov web space.

As a consequence, instead of the NASA Portal driving consolidation, it has fueled the dispersion problem, with www.nasa.gov becoming merely a single element in the overall ecosystem.

Where We Are Going

The last five years have seen a shift away from both Content Management System-centric and proprietary Enterprise Portal solutions to the use of Web Application Frameworks. Leading Internet companies such as Microsoft, Yahoo!, Amazon.com, Salesforce.com, and recently Google have all started building out web application frameworks, and web-services. These "cloud services" provide access to very low cost storage, network connectivity, and processing. These so-called "Software as a Service" and "Platform as a Service" models are revolutionizing the way both businesses and consumers are using the web as it relieves them of having to set up their own infrastructure while maintaining a flexible environment on which to build their businesses and web presence.

The next 50 years of space exploration will see an even greater wealth of information and content returning from space as instrument capabilities evolve and the number of satellites and missions increase. NASA will become more Internet-centric in its operations as more consumers of the information it provides expect to access this information on the web. The web as platform will become a greater strategic asset to the Agency as web technologies evolve, continuously providing new functionality and opportunities in many fields relevant to NASA's mission.

At the same time, the web is increasingly becoming the universal medium through which the American citizen interacts with the world. As a result, the public will have growing expectations concerning how they can interact with NASA's activities and missions, especially among the younger generation. These elevated expectations include real-time, or near real-time, insight into mission development and progress and an the ability to interact with NASA's information, people, systems, instruments, and robots.

> *"NASA will capture the imagination of the next generation of explorers on the web."*

Today's generation expects to *participate* in NASA's mission of exploration. For the public, this means flying through space with satellites, exploring other planets, and watching high definition video over the Internet of astronauts building a base on the moon. For innovators, scientists, and business, this means tapping into zettabytes of structured data on the web in real time.

The following table illustrates some of the Policy, Cultural, and Technology changes that will benefit NASA the faster they occur:

		Today	Future
Policy	Control	1. Editorial Board has Control 2. Control implemented by ownership of hardware 3. CIO and PAO write policy 4. Policy enforced on 1% of content	1. Community has control 2. Control implemented logically 3. Web Council writes policy 4. Policy enforced on 100% of content
	Funding	1. Funded at project (lowest) level	1. Funded at agency (highest) level
	Procurement	1. Tens of Millions of Dollars 2. Default to Procure 3. High level (integrators)	1. Millions of Dollars 2. Default to Partner (SAA) 3. Low level (providers)
Culture		1. Outsourced development 2. Outsourced content production 3. Loose community 4. Few content creators 5. Keyword competition	1. Open source development 2. Community content production 3. Strong community 4. Many content creators 5. Keyword collaboration
Technology	Software	1. Content Management System 2. Proprietary	1. Web Application Framework 2. Open Source
	Hardware	1. Physical Servers 2. Physically distributed	1. Virtual Servers 2. Physically consolidated(1)
	Distribution	1. Terabytes of Data 2. Inconsistent architecture 3. Inconsistent infrastructure 4. Local Storage (SAN)	1. Zetabytes of Data 2. Consistent Architecture 3. Consistent infrastructure 4. Storage Services (S3)

Figure 2. NASA's Web Presence Drivers: Today vs. Future.

Web Strategy

NASA's Web Strategy should include immediate web infrastructure consolidation into a single facility while simultaneously developing an agile, extensible, and flexible web application framework built on www.nasa.gov.

Deploying a modular and open source infrastructure will allow for future development and expanded functionality as the Internet continues to evolve and as user expectations and information access demands change.

The following objectives must be met by NASA's web strategy:

- Must make majority of NASA information easily discoverable, accessible, browsable, searchable, and useable for the public.
- Must provide flexible framework to support all NASA content creators (customers) both in functionality and infrastructure.
- Must enable every NASA employee to contribute content while maintaining quality and messaging standards.

- Must provide an agile infrastructure and architecture that can easily be aligned with rapidly evolving web technologies and Agency priorities.
- Must leverage NASA's unique position to establish hosting partnerships with cloud services for content syndication and distribution designed to offload traffic from www.nasa.gov and thereby provide significant cost savings to the Agency.
- Must provide increased security, availability, and accessibility.
- Must dramatically reduce the overall cost of maintaining NASA's web sites.
- Must consolidate architecture and infrastructure.
- Must provide increased visibility and coherency of NASA's activities and information by providing a bona-fide source of information.

Customer adoption is critical given NASA's cultural and policy environment. NASA's web strategy must support all content owners and provide a customer-focused approach such that content owners are willing to migrate to the new infrastructure, instead of being forced into it. The following requirements are key in achieving user adoption:

- **Content Ownership**. The new www.nasa.gov will allow content producers and owners – which should include, to some degree, all 65,000 NASA civil servants and contractors – to produce content. Policies will be developed to allow individuals, groups, or complete mission teams to set up and host content on the framework and retain ownership of their section of www.nasa.gov.

- **Dynamic Web Application Environment.** Advanced content producers must be able to go far beyond static content to deliver information and engage the public. The new www.nasa.gov must enable content producers to develop reusable, rich, dynamic, data-driven, and interactive applications.

- **Extreme Cost Savings.** The direct cost to the customer of building out a dynamic web site and hosting information (in terms of both storage and bandwidth) on www.nasa.gov must be at least an *order of magnitude less than* the cost of building infrastructure out independently. Today the cost is an *order of magnitude more than* building out infrastructure independently.

Implementation Plan

In order to achieve this vision, the policy, cultural, and technological changes described above and outlined in Figure 2 must be addressed. We propose a three-pronged approach:

1. Ratify the NASA Web Strategy Council (to effect policy changes)
2. Establish the NASA Operations Center (a new web operating paradigm)
3. Create a NASA Web Application Framework (the technology substrate)

1. NASA Web Strategy Council

The Web Strategy Council is the governance body responsible for interpreting federal web policies, implementing NASA web policies, and setting the overall direction for NASA's web

strategy. The NASA Web Strategy Council consists of web-oriented representatives and maps to the NASA governance model. The Web Strategy Council is the decision-making body for NASA's Public Web Presence.

A pilot of the Web Strategy Council has recently been launched and is providing a valuable forum for policy-oriented consensus building across the Agency. This body is charged with determining NASA's strategic direction by assessing the Agency's needs and driving technology funding and procurement decisions.

Once ratified, the NASA Web Strategy Council will:

1. Draft and ratify NASA web policies.
2. Endorse and support NASA's web strategy.
3. Ensure Agency, Mission Directorate, and Center needs are met through NASA's web strategy.
4. Solicit and incorporate feedback from key groups comprised of industry partners, universities, non-profit corporations, and other government agencies into the web strategy.
5. Facilitate communication among NASA web developers and content providers Agency-wide.

2.   NASA Web Operations Center

The Web Operations Center is responsible for the implementation, constant evolution, and daily operation of NASA's Public web infrastructure. The NASA Web Operations Center will leverage modern datacenter approaches, such as virtualization, to rapidly consolidate the Agency's infrastructure and markedly increase efficiency.

Once operational, the NASA Web Operations Center will:

1. Develop and maintain the NASA Web Application Framework (see below).
2. Manage the rapid consolidation of websites into the Web Operations Center virtual server environment to be located at a NASA Enterprise Data Center.
3. Provide real-time metrics and feedback from NASA's public web environment to assist the Web Strategy Council in key policy decisions.
4. Build a community of developers that will contribute to a repository of reusable content and functional modules that run on the NASA Web Application Framework.
5. Coordinate with Agency and Federal Information Architecture and taxonomy efforts.
6. Provide templates, modules, training, and other resources such that content producers and web developers can easily and effectively leverage these services.
7. Setup, promote, and manage content distribution and syndication collaborations with external parties such as Amazon, Google, Internet Archive, Microsoft, Sun, and Yahoo!.
8. Leverage NASA's content, information, and brand to serve the needs of its stakeholders, audiences, and communities.
9. Ensure conformance with NASA's Enterprise Architecture.
10. Ensure adherence to NASA IT security requirements for all websites and servers.

3. NASA Web Application Framework

The NASA Web Application Framework will empower users to build their own dynamic web content on www.nasa.gov without consideration for the underlying infrastructure. End-users (any of NASA's 65,000 civil servants and contractors) will be able to create content within the site with user-friendly, end-user configurable page templates and functional modules such as text/content, photo, blog, and wiki modules. If these "out of the box" modules do not meet user needs, these users will be able to write code leveraging the underlying NASA Application Programming Interface (API). An example of this approach is Force.com, or the Google Application Engine which launched the week this white paper was written.

Ultimately, both the content and functionality of the www.nasa.gov platform will be defined by its users. The agility of the platform will allow a wide variety of technologies to be adopted as they evolve and the whole system will be easily reconfigurable to incorporate these new technologies across the entire NASA domain.
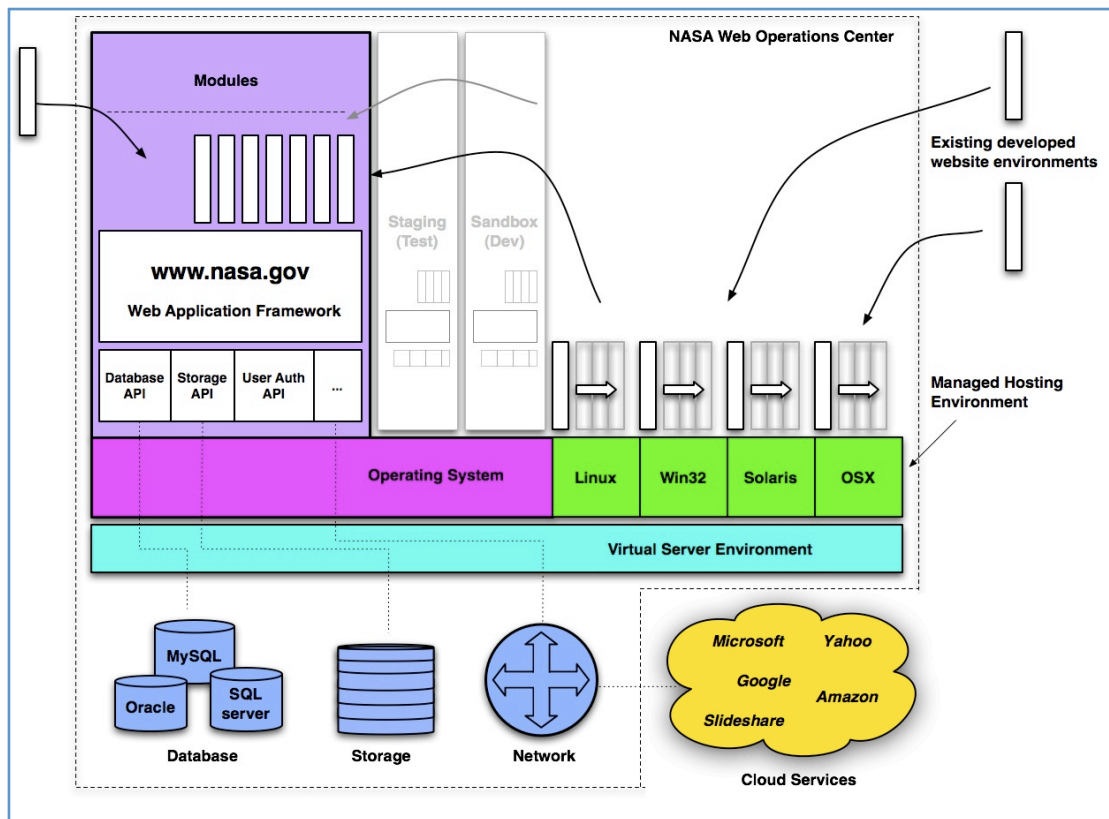

Figure 3. Proposed Architecture of NASA Web Operations Center and Web Application Framework

The proposed NASA architecture depicted in Figure 4 is consistent with modern Application Service Providers. The Web Application Framework will provide the APIs (building blocks) for NASA web developers to build future applications and will save countless hours in duplicated efforts across servers. This architecture will allow for scalability and economies of scale.